Defining Global Benchmarks in Bariatric Surgery: A Retrospective Multicenter Analysis of Minimally Invasive Roux-en-Y Gastric Bypass and Sleeve Gastrectomy

Gero, Daniel ; Raptis, Dimitri A ; Vleeschouwers, Wouter ; van Veldhuisen, Sophie L ; et al ; Hosa, Hanna ; Schmidt, Henner ; Bueter, Marco ; Lüthy, S G ; Puhan, Milo Alan

Abstract: OBJECTIVE To define "best possible" outcomes for bariatric surgery (BS)(Roux-en-Y gastric bypass [RYGB] and sleeve gastrectomy [SG]). BACKGROUND Reference values for optimal surgical outcomes in well-defined low-risk bariatric patients have not been established so far. Consequently, outcome comparison across centers and over time is impeded by heterogeneity in case-mix. METHODS Out of 39,424 elective BS performed in 19 high-volume academic centers from 3 continents between June 2012 and May 2017, we identified 4120 RYGB and 1457 SG low-risk cases defined by absence of previous abdominal surgery, concomitant procedures, diabetes mellitus, sleep apnea, cardiopathy, renal insufficiency, inflammatory bowel disease, immunosuppression, anticoagulation, BMI>50 kg/m and age>65 years. We chose clinically relevant endpoints covering the intra- and postoperative course. Complications were graded by severity using the comprehensive complication index. Benchmark values were defined as the 75th percentile of the participating centers' median values for respective quality indicators. RESULTS Patients were mainly females (78%), aged 38±11 years, with a baseline BMI 40.8 ± 5.8 kg/m. Over 90 days, 7.2% of RYGB and 6.2% of SG patients presented at least 1 complication and no patients died (mortality in nonbenchmark cases: 0.06%). The most frequent reasons for readmission after 90-days following both procedures were symptomatic cholelithiasis and abdominal pain of unknown origin. Benchmark values for both RYGB and SG at 90-days postoperatively were 5.5% Clavien-Dindo grade IIIa complication rate, 5.5% readmission rate, and comprehensive complication index 33.73 in the subgroup of patients presenting at least 1 grade II complication. CONCLUSION Benchmark cutoffs targeting perioperative outcomes in BS offer a new tool in surgical quality-metrics and may be implemented in quality-improvement cycle.ClinicalTrials.gov Identifier NCT03440138.

DOI: https://doi.org/10.1097/SLA.0000000000003512

Defining Global Benchmarks in Bariatric Surgery

A Retrospective Multicenter Analysis of Minimally Invasive Roux-en-Y Gastric Bypass and Sleeve Gastrectomy

Daniel Gero, MD,[*] Dimitri A. Raptis, MD, MSc, PhD,[†] Wouter Vleeschouwers, MD,[‡] Sophie L. van Veldhuisen, MD,[§] Andres San Martin, MD,[¶] Yao Xiao, MD,[||**] Manoela Galvao, MD,[††] Marcoandrea Giorgi, MD,[‡‡] Marine Benois, MD,[§§] Felipe Espinoza, MD,[¶¶] Marianne Hollyman, MD, PhD,[||||] Aaron Lloyd, MPH,[***] Hanna Hosa, MD,[*] Henner Schmidt, MD,[*] José Luis Garcia-Galocha, MD,[†††] Simon van de Vrande, MD,[‡‡‡] Sonja Chiappetta, MD,[§§§] Emanuele Lo Menzo, MD,[¶¶¶] Cristina Mamédio Aboud, RN, MSc,[||||||] Sandra Gagliardo Lüthy,[****] Philippa Orchard, MD,[††††] Steffi Rothe, MBA,[‡‡‡‡] Gerhard Prager, MD,[‡‡‡‡] Dimitri J. Pournaras, MD, PhD,[††††] Ricardo Cohen, MD,[||||||] Raul Rosenthal, MD,[¶¶¶] Rudolf Weiner, MD,[§§§] Jacques Himpens, MD, PhD,[‡‡‡§§§§] Antonio Torres, MD, PhD,[†††] Kelvin Higa, MD,[***] Richard Welbourn, MD,[||||] Marcos Berry, MD,[¶¶] Camilo Boza, MD,[¶¶] Antonio Iannelli, MD,[§§] Sivamainthan Vithiananthan, MD,[‡‡] Almino Ramos, MD,[††] Torsten Olbers, MD, PhD,[**¶¶¶¶] Matias Sepúlveda, MD,[¶] Eric J. Hazebroek, MD, PhD,[§] Bruno Dillemans, MD,[‡] Roxane D. Staiger, MD,[*] Milo A. Puhan, MD, PhD,[||||||||] Ralph Peterli, MD,[****] and Marco Bueter, MD, PhD[*]✉*

Objective: To define "best possible" outcomes for bariatric surgery (BS)(Roux-en-Y gastric bypass [RYGB] and sleeve gastrectomy [SG]).

Background: Reference values for optimal surgical outcomes in well-defined low-risk bariatric patients have not been established so far. Consequently, outcome comparison across centers and over time is impeded by heterogeneity in case-mix.

Methods: Out of 39,424 elective BS performed in 19 high-volume academic centers from 3 continents between June 2012 and May 2017, we identified 4120 RYGB and 1457 SG low-risk cases defined by absence of previous abdominal surgery, concomitant procedures, diabetes mellitus, sleep apnea, cardiopathy, renal insufficiency, inflammatory bowel disease, immunosuppression, anticoagulation, BMI>50 kg/m^2 and age>65 years. We chose clinically relevant endpoints covering the intra- and postoperative course. Complications were graded by severity using the comprehensive complication index. Benchmark values were defined as the 75th percentile of the participating centers' median values for respective quality indicators.

Results: Patients were mainly females (78%), aged 38±11 years, with a baseline BMI 40.8 ± 5.8 kg/m^2. Over 90 days, 7.2% of RYGB and 6.2% of SG patients presented at least 1 complication and no patients died (mortality in nonbenchmark cases: 0.06%). The most frequent reasons for readmission after 90-days following both procedures were symptomatic cholelithiasis and abdominal pain of unknown origin. Benchmark values for both RYGB and SG at 90-days postoperatively were 5.5% Clavien-Dindo grade ≥IIIa complication rate, 5.5% readmission rate, and comprehensive complication index ≤33.73 in the subgroup of patients presenting at least 1 grade ≥II complication.

Conclusion: Benchmark cutoffs targeting perioperative outcomes in BS offer a new tool in surgical quality-metrics and may be implemented in quality-improvement cycle.

ClinicalTrials.gov Identifier NCT03440138

Keywords: bariatric surgery, benchmark, complication, morbidity, outcome research, quality assessment, Roux-en-Y gastric bypass, sleeve gastrectomy

(*Ann Surg* 2019;270:859–867)

With growing complexity and cost of modern surgical practice, structured quality assessment became mandatory.[1,2] Benchmarking is among the most popular quality management tools in companies' process improvement cycles.[3] Benchmarking is a market-based learning method by which a company seeks to identify best practices that produce superior results in other firms, and to enhance its own competitive advantage by adopting them.[4] In the surgical literature procedure-specific outcome benchmarks are largely lacking.[5]

From the [*]Department of Surgery and Transplantation, University Hospital Zurich, Zurich, Switzerland; †Department of Hepatobiliary and Pancreas Surgery and Liver Transplantation, Royal Free Hospital, London, UK; ‡Department of General Surgery, AZ Sint Jan Brugge-Oostende, Brugge, Belgium; §Department of Surgery, Rijnstate Hospital/Vitalys Clinics, Arnhem, The Netherlands; ¶Bariatric and Metabolic Surgery Center, Dipreca Hospital, Las Condes, Santiago, Chile; ||Department of Surgery, Varberg Hospital, Varberg, Sweden; [**]Department of Surgery, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden; ††GastroObeso-Center—Advanced Institute In Bariatric And Metabolic Surgery, Sao Paulo, Brazil; ‡‡Department of Surgery, Alpert Medical School of Brown University/The Miriam Hospital, Providence, RI; §§Digestive Surgery and Liver Transplantation Unit, University Hospital of Nice, University Côte d'Azur, Nice, France; ¶¶Bariatric and Metabolic Center, Department of Surgery, Clinica Las Condes, Las Condes, Santiago, Chile; ||||Department of Upper Gastrointestinal and Bariatric Surgery, Musgrove Park Hospital, Taunton, UK; [***]Minimally Invasive and Bariatric Surgery, Fresno Heart and Surgical Hospital, Fresno, CA; †††Department of Surgery, Hospital Clínico San Carlos, Complutense University of Madrid, Madrid, Spain; ‡‡‡Department of General Surgery, AZ Sint-Blasius Hospital, Dendermonde, Belgium; §§§Department of Obesity and Metabolic Surgery, Sana Klinikum Offenbach, Offenbach, Germany; ¶¶¶The Bariatric and Metabolic Institute, Cleveland Clinic Florida, Weston, FL; ||||||Center for the treatment of Obesity and Diabetes - COD, Oswaldo Cruz German Hospital, Sao Paulo, Brazil; [****]Department of Visceral Surgery, Clarunis: St.Clara Hosptital and University Hospital Basel, Basel, Switzerland; ††††North Bristol Centre for Weight Loss Metabolic & Bariatric Surgery Southmead Hospital, Bristol, UK; ‡‡‡‡Department of Surgery, Vienna Medical University, Vienna, Austria; §§§§The European School of Laparoscopic Surgery, St Pierre University Hospital, Brussels, Belgium; ¶¶¶¶Linköping University, Institute for Clinical and Experimental Medicine, Department of Surgery Vrinnevi, Norrköping, Sweden; and ||||||||Epidemiology, Biostatistics and Prevention Institute, University of Zurich, Zurich, Switzerland.

✉marco.bueter@usz.ch.

R.P. and M.B. contributed equally.

TABLE 1. Criteria Used to Identify Participating Centers and ''Benchmark'' Cases

Center Inclusion Criteria	Patient Inclusion Criteria	Patient Exclusion Criteria
Annual caseload $\geq$ 200 bariatric cases, at least during the last year of the study period[19,20]	Age 18–65 yrs[21–23]	Previous intra-abdominal surgery (including previous bariatric surgery)[24,25]
Minimum 30 benchmark cases over the 5-yr study period for inclusion in the procedure-specific (RYGB or SG) establishment of global benchmarks	Low risk profile (please read ''exclusion criteria'')	Cardiovascular disease (eg, cardiac arrhythmia, stroke, coronary artery disease)[22]
Available prospective bariatric database	Preoperative BMI $\leq$ 50 kg/m^2 [26,27]	History of thromboembolic events and/or therapeutic anticoagulation[27]
Interest in bariatric outcomes, documented by $\geq$1 publication(s) on bariatric surgery	Laparoscopic primary Roux-en-Y gastric bypass or sleeve gastrectomy[24]	Diabetes mellitus (Type 1 and Type 2, as defined by the American Diabetes Association)[28,29]
''Clinical excellence'' or national reference centers[18]	Documented follow-up of at least 90 d[30]	Obstructive sleep apnea (recurrent episodes of upper airway collapse during sleep)[26,27]
	American Society of Anesthesiologists (ASA) score $<$ IV[31]	Chronic obstructive pulmonary disease (FEV1/FVC$<$0.7)[24]
		Chronic kidney disease (eGFR $<$30 mL/min/1.72 m^2)[22]
		Inflammatory bowel disease (ulcerative colitis, Crohn disease)[32]
		Immunosuppression therapy (ie,. steroids, calcineurin inhibitors, etc.)[33,34]
		Associated surgical procedures (ie,. cholecystectomy, hiatoplasty, liver biopsy)[17]

Recent studies in the field of visceral surgery established benchmark cutoffs for best achievable patient-centered outcomes in well-defined low-risk patient cohorts, allowing comparison among centers and patient groups over time and between procedures.[2,6–9]

Bariatric surgery (BS) remains the most effective treatment for severe obesity and associated diseases.[10] The annual caseload of BS worldwide has doubled during the past decade, and approached 700,000 operations in 2016.[11] Together, the Roux-en-Y gastric bypass (RYGB) and the sleeve gastrectomy (SG) constitute more than 80% of bariatric procedures worldwide.[11] Although frequently performed, both procedures and follow-up care are not highly standardized,[12,13] posing a challenge in defining evidence-based cutoffs for quality indicators for the peri- and postoperative course.[14]

Our aim was to identify the highest achievable quality (ie, the global benchmarks) in BS, by assessing patient-centered outcome indicators in low-risk patients who underwent SG or RYGB in high-volume bariatric centers. The identified benchmarks are expected to improve surgical quality by providing ''goals'' in postoperative outcomes and may therefore assist patients and healthcare providers in medical decision-making.

METHODS

Study Design

The establishment of benchmarks in BS followed a standardized methodology, previously applied in visceral surgery.[2,6–8,15] We performed a multicentric retrospective cohort study based on prospective institutional databases to define best achievable surgical outcomes in primary laparoscopic RYGB and SG.

First, a large patient cohort from international expert centers was gathered via personal invitation of distinguished surgeons. Eligible centers had to meet the criteria listed in Table 1.[16–33] The final collaborative consortium included 19 centers: 12 from Europe (Arnhem, Basel, Brussels/Dendermonde, Bristol, Bruges, Gothenburg, Madrid, Nice, Offenbach, Taunton, Wien, and Zurich), 3 from USA (Fresno, Providence, Weston), and 4 from South America (Santiago de Chile, São Paulo [2 centers in each city]).

Second, to define the ''benchmark bariatric patient,'' evidence-based criteria associated with a lower postoperative complication rate were applied (Table 1). Each center had to include all consecutive benchmark cases with a documented follow-up of minimum 90-days (including mortality), operated over a 5-year period (June 1, 2012 to May 31, 2017).

Third, relevant outcome indicators for surgical quality were assessed. To adjust for variability, median values of continuous variables and the proportions of categorical variables were calculated for each participating center. Benchmark cutoffs, indicating ''best achievable'' result for each outcome indicator, were set at the 75th percentile of the centers' median values. Additionally, range and median (ie, ''premium result'') of each indicator was computed. The study protocol was preregistered on ClinicalTrials.gov (NCT03440138). The approval of the Cantonal Ethics Committee of Zurich, as well as of the ethical board of each respective center, was obtained before data analysis.

Outcome Variables of Interest

Local investigators retrieved patient-specific data and uploaded them to a secure and anonymized online data-entry management system

This study was entirely funded by the institutions of the coauthors, without any external support.
The study was presented as oral presentation at the 2019 meeting of the European Surgical Association in Madrid, Spain, on May 17, 2019.
Disclosures: Prof. R.P. reports consulting fees from Johnson & Johnson, outside the submitted work. Prof. M.B. reports personal fees from Johnson & Johnson and Medtronic, outside the submitted work. Dr. R.W. reports nonfinancial support for attending conferences from Ethicon Endo-Surgery, personal fees from consultancy fees from Novo Nordisk, outside the submitted work. Dr. R.C. reports personal fees from GI Dynamics, grants from Johnson & Johnson Medical, Brazil, outside the submitted work. Prof. J.H. reports personal fees from Ethicon, personal fees from Medtronic, outside the submitted work. Dr. B.D. reports consulting fees from Covidien Inc. and Johnson & Johnson, outside the submitted work. Dr. M.B. reports an educational (fellowship) grant from Medtronic. Prof E.J.H. reports consulting fees from Johnson & Johnson, outside the submitted work. The remaining authors declare no conflicts of interest in association with the present study.
The Cantonal Ethics Committee of Zurich approved this study (BASEC-Nr. 2017-01652). The study was preregistered at ClinicalTrials.gov (Identifier: NCT03440138).
Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.annalsofsurgery.com).

ISSN: 0003-4932/19/27005-0860
DOI: 10.1097/SLA.0000000000003512

provided by the University Hospital Zurich.[34] Data were audited and checked for completeness by DG and included baseline characteristics of patients (age, sex, body-mass index [BMI], comorbidities), operation characteristics, inpatient complications by severity according to the Clavien–Dindo (CD) grading system,[35,36] length of stay, readmissions (time from operation, reason, and treatment), last follow-up, and yearly postoperative BMI. To enable the assessment of cumulative morbidity over time, the Comprehensive Complication Index (CCI) was used.[37] The CCI expresses morbidity on a continuous numeric scale from 0 (no complications) to 100 (death) by weighing all postoperative complications according to the CD classification. Relevant bariatric complications, such as staple line/anastomotic leak, anastomotic or gastric tube stenosis, internal hernia, marginal ulcer at the gastrojejunostomy, were additionally analyzed.[16,38,39] Postoperative weight loss was expressed as %total weight loss, and excess body mass index loss (EBMIL) (% excess BMI loss, with BMI = 25 kg/m^2 considered as normal).[40] Benchmark cutoffs for overall inpatient costs were calculated separately for patients with or without any surgical complications by using the algorithm developed by Staiger et al.[41]

Proof of Concept

To validate the need for outcome benchmarks in BS, we applied 3 complementary measures. First, we collected data on 90-day mortality of all RYGB and SG cases operated during the study period. Second, we selected 1 participating center (Number 2 in Fig. 1) to provide outcome data of all nonbenchmark RYGB cases. Third, we identified previously published studies by the participating centers reporting 30-day major complications rates in consecutive and secondary RYGB cohorts to objectify the additional burden of "high-risk" cases on the early postoperative morbidity.

Statistical Analysis

Centers that contributed with <30 cases per type of BS (RYGB or SG) were excluded from analyses in the respective operation subgroup.[42] Discrete variables were described using count (percent), and continuous variables were described using medians (with interquartile range). Kaplan–Meier curve was used to describe the occurrence of postoperative CD >II complications over observation time. Statistical analysis was performed using the R software 3.5.1 (R Foundation, Vienna, Austria).

RESULTS

Out of all 39,424 consecutive elective BS cases (RYGB, SG, bilio-pancreatic diversion, gastric banding, single anastomosis gastric bypass, single anastomosis duodenoileal bypass, gastric plication) performed over 5 years in the 19 included centers, 4120 RYGB and 1457 SG benchmark cases were identified (Supplementary Figure 1, http://links.lww.com/SLA/B717). The proportion of benchmark cases within the case mix of participating centers varied from 4% to 69% (Fig. 1). Baseline characteristics of patients and procedures are presented in Supplementary Table 1, http://links.lww.com/SLA/B717. Majority of centers had a >90% uneventful postoperative course rate during the first 90-days for both procedures (Supplementary Figure 2, http://links.lww.com/SLA/B717). The cumulative hazard of CD grade >II events after BS was below 4% at 90 d (Fig. 2); nevertheless, it increased constantly over time during the first 2 postoperative years (Supplementary Figure 3, http://links.lww.com/SLA/B717). The most common reasons for readmissions until last follow-up are shown in Figure 3 and Supplementary Table 2, http://links.lww.com/SLA/B717. There was a great between-center variability in the size of benchmark cohorts, in median length of follow-up, and in the cumulative hazard of reinterventions beyond 90 days (Supplementary Figure 4, http://links.lww.com/SLA/B717). Centers with higher caseload showed a trend toward achieving lower mean CCI over 90 days; however, these correlations were not statistically

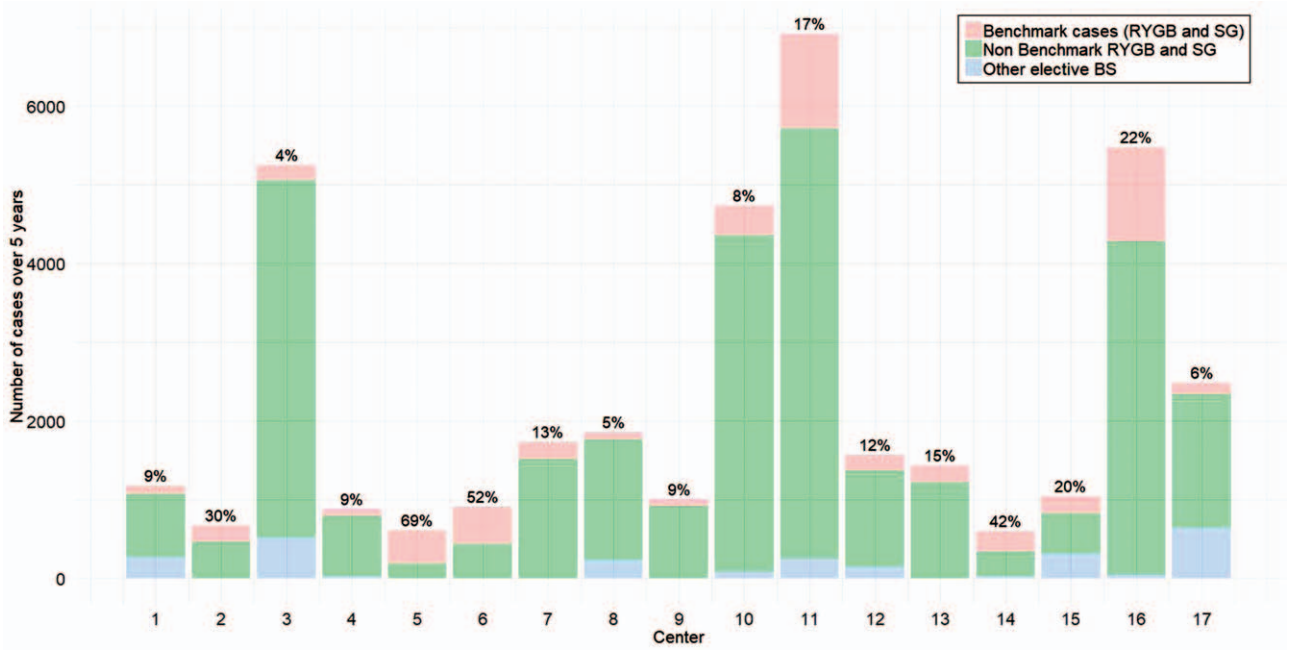

FIGURE 1. Case mix of elective bariatric surgery (BS) in participating centers over the 5-yr inclusion period (June 2012–May 2017). Percentages show the proportion of benchmark Roux-en-Y gastric bypass (RYGB) and sleeve gastrectomy (SG) cases within the total elective BS caseload.

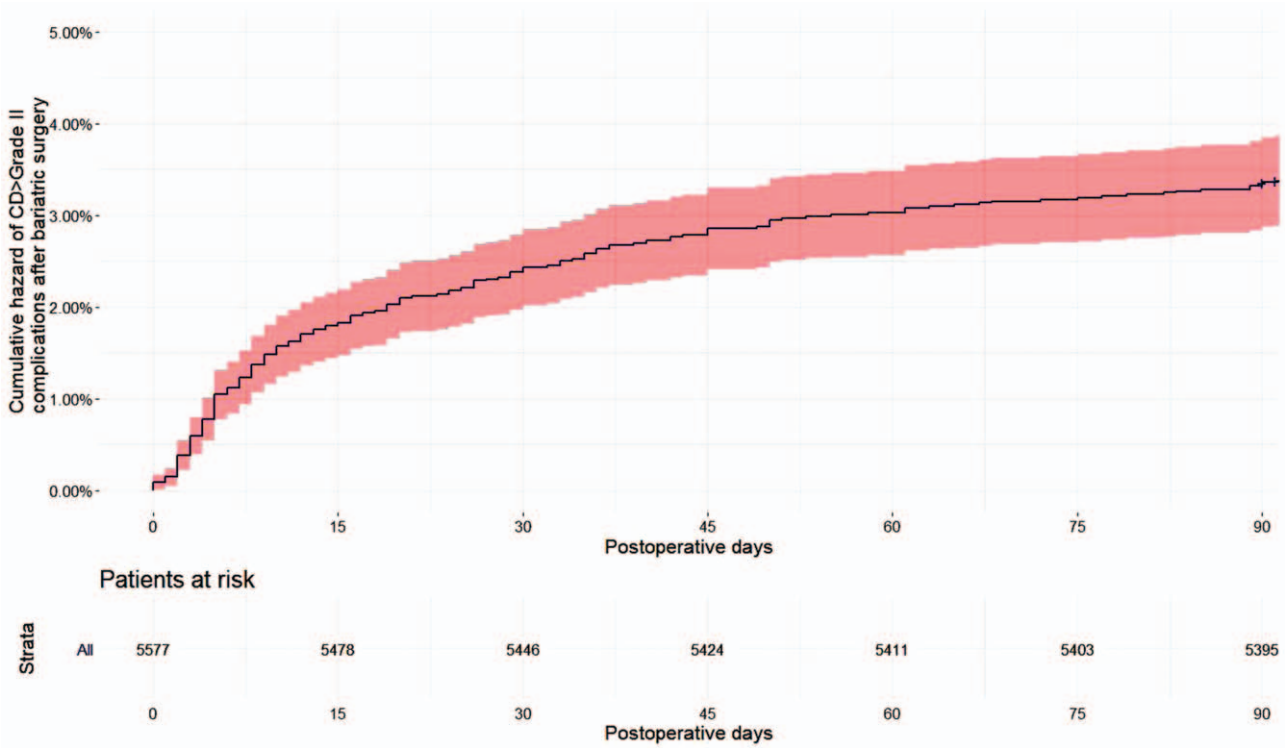

FIGURE 2. Cumulative hazard of Clavien–Dindo grade > II events in benchmark patients during the first 90-d after bariatric surgery (Roux-en-Y gastric bypass and sleeve gastrectomy).

FIGURE 3. Cumulative incidence (%) of the most common reasons for readmission in benchmark patients after bariatric surgery. A, Roux-en-Y gastric bypass (n = 4120, median follow-up = 1.9 yr, [range: 0.25–6 yrs]). B, Sleeve gastrectomy (n = 1457, median follow-up = 1.6 yr, [range: 0.25–6 yrs]). GERD indicates gastro-esophageal reflux disease.

TABLE 2. Benchmark Cutoffs for Roux-en-Y Gastric Bypass (75th Percentile of Centers' Median)

1. Perioperative Course			
Operation duration		≤2 h	
Conversion to open surgery		0%	
Intraoperative blood transfusions		0%	
Postoperative blood transfusions		≤2%	
Postoperative ICU admission		≤0.14%	
ICU stay in patients admitted to ICU		≤1 d	
Hospital stay		≤4 d	
Hospital cost in CH or USA/in the EU		16,203 CHF or USD / 5402 EUR	
Hospital cost in patients with complications in CH or USA / in the EU		26,485 CHF or USD / 8705 EUR	
2. Morbidity and Mortality	**Until Discharge**	**Until 30-d**	**Until 90-d**
Uneventful postoperative course	>94%	>91%	>90%
Readmission	—	≤4%	≤5.5%
Reoperation	≤2%	≤2.5%	≤4%
Any complication	≤6%	≤9%	≤10%
Complication grade ≥ IIIa	≤3.5%	≤5%	≤5.5%
Mortality	0%	0%	0%
CCI In patients with at least 1 Clavien–Dindo Grade ≥II complication	≤26.2	≤32.5	≤33.73
Signature complications			
Anastomotic leak	0%	≤1.1%	≤1.3%
Stenosis of the anastomosis	0%	≤0.3%	≤1.2%
Postoperative bleeding	≤2.2%	≤2.2%	≤2.2%
Small bowel obstruction/internal hernia	≤1.4%	≤2.1%	≤2.1%
Wound infection	≤0.5%	≤0.5%	≤0.5%
Marginal ulcer	0%	≤0.3%	≤1.5%

significant (Supplementary Figure 5, http://links.lww.com/SLA/B717).

Benchmark Cutoffs of Quality Indicators

Outcome benchmarks of RYGB and SG are shown in Tables 2 and 3, with additional data including range and median in Supplementary Table 3, http://links.lww.com/SLA/B717.

Roux-en-Y Gastric Bypass

At baseline, the cohort's mean age was 38.2 ± 11.1 years with a BMI of 41.3 ± 6.2 kg/m^2. Before discharge, 3.4% of patients presented at least 1 complication. Readmissions due to grade >II events occurred in 2.5%, 4.1%, 5.5%, and 9.4% of cases at postoperative days 30, 90, 180, and 365. Ninety-day and 1-year mortality were 0% and 0.02% (1 patient died from a cardiovascular event on

TABLE 3. Benchmark Cutoffs for Sleeve Gastrectomy (75th Percentile of Centers' Median)

1. Perioperative Course			
Operation duration		≤1.5 h	
Conversion to open surgery		0%	
Intraoperative blood transfusions		0%	
Postoperative blood transfusions		≤1.3%	
Postoperative ICU admission		0%	
ICU stay in patients admitted to ICU		≤4 d	
Hospital stay		≤3 d	
Hospital cost in CH or USA/in the EU		16,204 CHF or USD / 5402 EUR	
Hospital cost in patients with complications in CH or USA / in the EU		25,949 CHF or USD / 8650 EUR	
2. Morbidity and Mortality	**Until Discharge**	**Until 30-d**	**Until 90-d**
Uneventful postoperative course	>92%	>89%	>88%
Readmission	—	≤4%	≤5.5%
Reoperation	≤2%	≤2%	≤3%
Any complication	≤8%	≤11%	≤12%
Complication grade ≥ IIIa	≤2.5%	≤5%	≤5.5%
Mortality	0%	0%	0%
CCI In patients with at least 1 Clavien–Dindo Grade ≥II complication	≤26.22	≤32.53	≤33.73
Signature complications			
Staple line leak	0%	≤0.15%	≤0.15%
Dysphagia/Stenosis of the gastric tube	0%	≤0.14%	≤0.27%
Postoperative bleeding	≤1.7%	≤1.7%	≤1.7%
Small bowel obstruction	0%	0%	0%
Wound infection	0%	0%	0%

postoperative-day 211). At 1-year (follow-up: 82.5%), the cohort's mean BMI was $27.7 \pm 4\,kg/m^2$, EBMIL: $86.8 \pm 25.5\%$ and %weight loss: $32.7 \pm 8.3\%$.

Sleeve Gastrectomy

At baseline, the cohort's mean age was 37.0 ± 10.8 years with a BMI of $38.9 \pm 5.2\,kg/m^2$. Before discharge, 3.6% of patients presented at least 1 complication. Readmissions due to grade >II events occurred in 2.5%, 3.1%, 3.7%, and 5.9% of cases at postoperative days 30, 90, 180, and 365. One-year mortality was zero. At 1-year (follow-up: 68.2%), the cohort's mean BMI was $28 \pm 4.9\,kg/m^2$, EBMIL: $84.3 \pm 37.6\%$ and %weight loss: $27.5 \pm 10.2\%$.

Proof of Concept

The 90-day mortality rate in the nonbenchmark cases was 0.05% (11/21,830) after RYGB and 0.09% (8/8813) after SG. The 90-day postoperative outcomes of nonbenchmark RYGB patients operated at Center Number 2 (n = 468) are compared with the benchmark cutoffs for the same quality indicators in Supplementary Table 4, http://links.lww.com/SLA/B717. A literature search identified 10 publications covering 3993 RYGB cases operated at the participating centers (Supplementary Table 5, http://links.lww.com/SLA/B717), with a median 30-day major complication rate of 10.75%.

DISCUSSION

This multicenter study established outcome benchmarks for the 2 most frequently performed bariatric procedures by applying a recently developed standardized methodology.[2] In the current report of low-risk BS patients operated in 1 of 19 high-volume referral centers located on 3 continents, main findings were a zero 90-day mortality rate, a low early postoperative morbidity rate with the majority of reinterventions occurring after the first 90-days. The cohort's 1-year percentage weight loss was comparable to the mean procedure-specific pooled outcomes published in the 2018 IFSO global registry report.[43]

Identified outcome benchmarks may serve as a reference for bariatric centers to compare their own outcomes in similarly low risk or even all bariatric patients and to take action when eventual performance gaps are identified. So far, these attempts were impeded by great variability in case mix between centers.[17,43,44] The approach is reminiscent to the concept of propensity scoring study participants, where randomization at baseline is mimicked by cohorts that are comparable on main measured covariates.[45] In practice, random allocation of patients to different bariatric centers is not feasible; therefore, the concept of establishing global benchmarks based on multicentric data of low-risk patients seems to be an appealing alternative to allow comparison of outcomes and thus, of surgical quality.

This study aimed to represent the ''real world'' by including European, Northern and Southern American centers led by recognized experts in the field of BS. All centers had sufficient caseload, a prospective bariatric database and previous publication(s) on surgical outcomes. The selection of benchmark patients was performed by a strict and stepwise risk stratification aiming to identify the ''healthiest'' BS candidates with the least expected complications. Each submitted case was read by one of the principal investigators in Zurich, and clarification was requested from the coinvestigators in case of incompletely submitted case report forms. Our protocol focused on CD >II events, given that medically treated complications of BS are frequently managed by nonsurgeon healthcare providers outside of bariatric centers and thus, do not obviously appear in institutional databases.[46]

All included patients were operated in academic centers with teaching assignments, and consequently, with potentially increased operation duration.[47] This may explain why established outcome benchmarks for operation times in our study were not shorter, but in a similar range as those reported for all laparoscopic RYGB and SG (including high-risk patients and redo surgery) performed in the USA in 2015 to 2016.[45]

Length of stay is not a pure quality indicator, since it highly depends on health care systems and at some extent, on the attitude of patients. We are witnessing a trend toward shorter inpatient stays following BS. The safety of early discharge on the first postoperative day in a selective group of bariatric patients without significant comorbidities is supported by the 2015 dataset of the MBSAQIP.[48] Furthermore, SG in selected low-risk patients has been recently found to be safely feasible even as an outpatient procedure.[49] Therefore, it is of no surprise that the 2018 IFSO global registry report showed shorter lengths of stay for RYGB (2 vs 4 d) and a similar one for SG (3 d) in comparison with the benchmarks identified in the current study, which were derived from a cohort operated between 2012–2017.[43]

The study revealed an interesting pattern in postoperative morbidity, as most reoperations and reinterventions occurred beyond 90 days. This is in line with postbariatric readmission rates observed in The Danish National Health Surveys,[50] but it is in contrast to previous reports on the temporal occurrence of reinterventions after other types of major abdominal surgeries (ie, hepatectomy, esophagectomy), where the vast majority of CD grade >II events were recorded within 30 days.[6,7] Several explanations to this finding may be possible. First, both BS patients and bariatric centers are committed to perform a lifelong follow-up, which may not be the case for other types of surgeries. Second, severe obesity is a chronic disease, which is often characterized by cyclic episodes of weight loss and weight regain, as well as a higher risk for the development of a series of associated conditions that may require surgical care (ie, gallstone disease, GERD).[51] In this study, symptomatic cholelithiasis, GERD, and weight regain were among the most frequent reasons of long-term postbariatric readmissions, although the prevalence of these pathologies may not entirely depend on surgical performance at the index procedure, and increases with length of follow-up. Abdominal pain of unknown origin was the most common reason of readmission after RYGB and the second most frequent one after SG. This is somewhat surprising and may be in part related to the retrospective design of the study. Surgical databases often record the chief complaint of the patients' at presentation and are not always updated by the definitive diagnosis retained at the end of the often time-consuming work-up.

The concept of establishing benchmarks in BS was validated by complementary approaches. First, we found that compared with the zero 90-day mortality of benchmark patients, the same centers recorded a 0.06% mortality rate following RYGB and SG in nonbenchmark cases. This small difference emphasizes the need for quality indicators focusing on postoperative morbidity. Second, we found that in 1 participating center the overall morbidity in the nonbenchmark cases was above the global benchmark cutoff, mainly due to the higher frequency of CD ≤IIIb complications. Third, the previous outcome reports on consecutive or secondary RYGB published by the participating centers showed a higher 30-day major complication rate than the benchmark cutoff (10.75% vs 5%), highlighting the additional burden of postoperative morbidity observed in ''high-risk'' cases.

This study has some inherent limitations. *First*, the quality of current institutional surgical databases seems to be suboptimal for capturing the full spectrum of postoperative morbidity, especially beyond 90 days. This could be improved in the future by external auditing of surgical databases and by replacing self-reporting with automatized data input methods. Ibrahim et al[17] also found a wide variation in postoperative severe complications rates among

accredited BS centers in the USA. In the current study, it was not possible to judge whether between-center differences reflected variability in surgical performance or were due to missing data regarding postoperative events. Although benchmark cutoffs were established until 90-days postoperatively in patients with 100% follow-up, we cannot exclude that some complications may have been underreported. *Second*, to minimize the confounding effect of a potentially hostile intra-abdominal status and of associated procedures on postoperative morbidity, we excluded cases with previous abdominal surgery and with additional nonbariatric procedures performed in combination with the index procedure, including cholecystectomies. This may at least partly explain why symptomatic cholelithiasis ranked among the most frequent postbariatric reasons for readmission.[51] *Third*, the current methodology of global surgical benchmark establishment is bounded by logistic obstacles, leading to a considerable burden for its future reproductions. The rapid evolution of surgical and endoscopic bariatric procedures, as well as the increasing caseload and experience of referral centers will mandate the regular update of bariatric benchmarks. Ultimately, this process needs to be automatized by development and adaptation of BS registries. *Fourth,* the case mix in the presented study does not reflect current practices: SG, the dominant operation worldwide,[11] represented only 26% of benchmark cases, thus future studies should aim to update benchmarks for SG with the inclusion of centers with a higher experience with this technique.

In conclusion, we consider this project as an inaugural study introducing the concept of benchmarking in BS. The surgical community's genuine desire to improve patients' postoperative outcomes has a crucial role in increasing penetrance of BS and in decreasing complication-related patient discomfort and healthcare expenditures. The concept of benchmarking is expected to be embedded in surgical quality improvement cycles, and to stimulate the need for comprehensive large databases allowing precise and timely identification of both global benchmarks and institutional outcomes.

ACKNOWLEDGMENTS

First, the authors are indebted to Professor Pierre-Alain Clavien, MD, PhD for his substantial input and critical review of the manuscript. In addition, the authors are grateful to Laura Deden, TP, Maria Fonseca Mora, MD, Xavier Muller, MD, as well as to all the clinical nurses, data-managers (in particular to Mrs Lubach Iris Sung Eun) and residents for their valuable contribution to the data-collection process.

REFERENCES

1. Vonlanthen R, Lodge P, Barkun JS, et al. Toward a consensus on centralization in surgery. *Ann Surg*. 2018;268:712–724.
2. Staiger RD, Schwandt H, Puhan MA, et al. Improving surgical outcomes through benchmarking. *Br J Surg*. 2019;106:59–64.
3. Vorhies DW, Morgan NA. Benchmarking marketing capabilities for sustainable competitive advantage. *J Marketing*. 2005;69:80–94.
4. Camp RC. *Business Process Benchmarking: Finding and Implementing Best Practices.* Milwaukee, WI: ASQC Quality Press; 1995.
5. Shoar S, Mahmoudzadeh H, Naderan M, et al. Long-term outcome of bariatric surgery in morbidly obese adolescents: a systematic review and meta-analysis of 950 patients with a minimum of 3 years follow-up. *Obes Surg*. 2017;27:3110–3117.
6. Rossler F, Sapisochin G, Song G, et al. Defining benchmarks for major liver surgery: a multicenter analysis of 5202 living liver donors. *Ann Surg*. 2016;264:492–500.
7. Schmidt HM, Gisbertz SS, Moons J, et al. Defining benchmarks for transthoracic esophagectomy: a multicenter analysis of total minimally invasive esophagectomy in low risk patients. *Ann Surg*. 2017;266:814–821.
8. Muller X, Marcon F, Sapisochin G, et al. Defining benchmarks in liver transplantation: a multicenter outcome analysis determining best achievable results. *Ann Surg*. 2018;267:419–425.
9. Helminen O, Mrena J, Sihvo E. Benchmark values for transthoracic esophagectomy are not set as the defined ''best possible''—a validation study. *J Thorac Dis*. 2018;10:4085–4093.
10. Samczuk P, Luba M, Godzien J, et al. Gear mechanism of bariatric interventions revealed by untargeted metabolomics. *J Pharm Biomed Anal*. 2018;151:219–226.
11. Angrisani L, Santonicola A, Iovino P, et al. IFSO worldwide survey 2016: primary, endoluminal, and revisional procedures. *Obes Surg*. 2018;28:3783–3794.
12. Edholm D, Ottosson J, Sundbom M. Importance of pouch size in laparoscopic Roux-en-Y gastric bypass: a cohort study of 14,168 patients. *Surg Endosc*. 2016;30:2011–2015.
13. Varban OA, Niemann A, Stricklen A, et al. Far from standardized: using surgical videos to identify variation in technique for laparoscopic sleeve gastrectomy. *J Laparoendosc Adv Surg Tech A*. 2017;27:761–767.
14. Sabench Pereferrer F, Dominguez-Adame Lanuza E, Ibarzabal A, et al. Quality criteria in bariatric surgery: consensus review and recommendations of the Spanish Association of Surgeons and the Spanish Society of Bariatric Surgery. *Cir Esp*. 2017;95:4–16.
15. Sanchez-Velazquez P, Muller X, Malleo G, et al. Benchmarks in pancreatic surgery: a novel tool for unbiased outcome comparisons. *Ann Surg*. 2019;270:211–218.
16. Tustumi F, Bernardo WM, Santo MA, et al. Cholecystectomy in patients submitted to bariatric procedure: a systematic review and meta-analysis. *Obes Surg*. 2018;28:3312–3320.
17. Ibrahim AM, Ghaferi AA, Thumma JR, et al. Variation in outcomes at bariatric surgery centers of excellence. *JAMA Surg*. 2017;152:629–636.
18. The European Accreditation Council for Bariatric Surgery. IFSO-EC Accredited Centre of Excellence in Bariatric and Metabolic Surgery: Code of Practice 2018. Available at: https://www.eacbs.com/site/index.php/ifso-endorsed-coe-programs/centre-of-excellence-coeprogram/requirements. 2018. Accessed July 18, 2019.
19. Krell RW, Finks JF, English WJ, et al. Profiling hospitals on bariatric surgery quality: which outcomes are most reliable? *J Am Coll Surg*. 2014;219:725–734.
20. Al-Kurd A, Grinbaum R, Mordechay-Heyn T, et al. Outcomes of Sleeve Gastrectomy in Septuagenarians. *Obes Surg*. 2018;28:3895–3901.
21. Quilliot D, Sirveaux MA, Nomine-Criqui C, et al. Evaluation of risk factors for complications after bariatric surgery. *J Visc Surg*. 2018;155:201–210.
22. Buchwald H, Estok R, Fahrbach K, et al. Trends in mortality in bariatric surgery: a systematic review and meta-analysis. *Surgery.* 2007;142:621–632.
23. Husain F, Jeong IH, Spight D, et al. Risk factors for early postoperative complications after bariatric surgery. *Ann Surg Treat Res*. 2018;95:100–110.
24. Benois M, Sebastianelli L, Morisot A, et al. Revisional but not conversional gastric bypass surgery increases the risk of leaks: review of 176 redo out of 932 consecutive cases. *Obes Surg*. 2018;28:2903–2911.
25. Fernandez AZ Jr, DeMaria EJ, Tichansky DS, et al. Experience with over 3,000 open and laparoscopic bariatric procedures: multivariate analysis of factors related to leak and resultant mortality. *Surg Endosc*. 2004;18:193–197.
26. Longitudinal Assessment of Bariatric Surgery C, Flum DR, Belle SH, et al. Perioperative safety in the longitudinal assessment of bariatric surgery. *N Engl J Med.* 2009;361:445–454.
27. Kabir A, Mousavi S, Pazouki A. The complications of bariatric surgery patients with type 2 diabetes in the world: a systematic review and meta-analysis. *Curr Diabetes Rev.* 2019;15:49–61.
28. Kumar R, Nandhini LP, Kamalanathan S, et al. Evidence for current diagnostic criteria of diabetes mellitus. *World J Diabetes*. 2016;7:396–405.
29. Melissas J, Stavroulakis K, Tzikoulis V, et al. Sleeve Gastrectomy vs Roux-en-Y Gastric Bypass. Data from IFSO-European Chapter Center of Excellence Program. *Obes Surg*. 2017;27:847–855.
30. ASA. ASA PHYSICAL STATUS CLASSIFICATION SYSTEM 2014. Available at: https://www.asahq.org/resources/clinical-information/asa-physical-status-classification-system. 2018. Accessed July 18, 2019.
31. Shoar S, Shahabuddin Hoseini S, Naderan M, et al. Bariatric surgery in morbidly obese patients with inflammatory bowel disease: a systematic review. *Surg Obes Relat Dis*. 2017;13:652–659.
32. Moulla Y, Lyros O, Bluher M, et al. Feasibility and safety of bariatric surgery in high-risk patients: a single-center experience. *J Obes*. 2018;2018:7498258.
33. Andalib A, Aminian A, Khorgami Z, et al. Early postoperative outcomes of primary bariatric surgery in patients on chronic steroid or immunosuppressive therapy. *Obes Surg*. 2016;26:1479–1486.
34. Raptis DA, Mettler T, Fischer MA, et al. Managing multicentre clinical trials with open source. *Inform Health Soc Care*. 2014;39:67–80.
35. Dindo D, Demartines N, Clavien PA. Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. *Ann Surg*. 2004;240:205–213.

36. Clavien PA, Barkun J, de Oliveira ML, et al. The Clavien-Dindo classification of surgical complications: five-year experience. *Ann Surg*. 2009;250:187–196.
37. Slankamenac K, Graf R, Barkun J, et al. The comprehensive complication index: a novel continuous scale to measure surgical morbidity. *Ann Surg*. 2013;258:1–7.
38. Poelemeijer YQM, Liem RSL, Vage V, et al. Perioperative outcomes of primary bariatric surgery in north-western Europe: a pooled multinational registry analysis. *Obes Surg*. 2018;28:3916–3922.
39. Duvoisin C, Favre L, Allemann P, et al. Roux-en-Y gastric bypass: ten-year results in a cohort of 658 patients. *Ann Surg*. 2018;268:1019–1025.
40. van de Laar A. Bariatric Outcomes Longitudinal Database (BOLD) suggests excess weight loss and excess BMI loss to be inappropriate outcome measures, demonstrating better alternatives. *Obes Surg*. 2012;22:1843–1847.
41. Staiger RD, Cimino M, Javed A, et al. The Comprehensive Complication Index (CCI(R)) is a novel cost assessment tool for surgical procedures. *Ann Surg*. 2018;268:784–791.
42. Martinez-Abrain A. Is the 'n=30 rule of thumb' of ecological field studies reliable? A call for greater attention to the variability in our data. *Animal Biodiversity Conservation*. 2014;37:95–100–.
43. Welbourn R, Hollyman M, Kinsman R, et al. Bariatric Surgery Worldwide: Baseline Demographic Description and One-Year Outcomes from the Fourth IFSO Global Registry Report 2018. *Obes Surg*. 2019;29:782–795.
44. Poelemeijer YQM, Marang-van de Mheen PJ, Wouters M, et al. Textbook outcome: an ordered composite measure for quality of bariatric surgery. *Obes Surg*. 2019;29:1287–1294.
45. Sebastian R, Howell MH, Chang KH, et al. Robot-assisted versus laparoscopic Roux-en-Y gastric bypass and sleeve gastrectomy: a propensity score-matched comparative analysis using the 2015-2016 MBSAQIP database. *Surg Endosc*. 2019;33:1600–1612.
46. Slankamenac K, Nederlof N, Pessaux P, et al. The comprehensive complication index: a novel and more sensitive endpoint for assessing outcome and reducing sample size in randomized controlled trials. *Ann Surg*. 2014;260:757–762.
47. Major P, Wysocki M, Dworak J, et al. Are bariatric operations performed by residents safe and efficient? *Surg Obes Relat Dis*. 2017;13:614–621.
48. Ardila-Gatas J, Sharma G, Lloyd SJ-A, et al. A nationwide safety analysis of discharge on the first postoperative day after bariatric surgery in selected patients. *Obes Surg*. 2019;29:15–22.
49. Surve A, Cottam D, Zaveri H, et al. Does the future of laparoscopic sleeve gastrectomy lie in the outpatient surgery center? A retrospective study of the safety of 3162 outpatient sleeve gastrectomies. *Surg Obes Relat Dis*. 2018;14:1442–1447.
50. Neermark S, Tolstrup JS, Becker PU, et al. Increased gastrointestinal surgical hospital burden after laparoscopic gastric bypass for obesity: a nationwide 5-years follow-up study. *Ann Surg* 9000; Publish Ahead of Print.
51. Hamdan K, Somers S, Chand M. Management of late postoperative complications of bariatric surgery. *Br J Surg*. 2011;98:1345–1355.

DISCUSSANTS

Nicolò de Manzini (Trieste, Italy):

This study is well structured using an original method based on economic models. The idea of finding some outcome benchmarks for ''easy'' cases of bariatric surgery, in absence of international recognized data, could certainly be useful in future comparisons.

However, the benchmark group represents only 14.1% of the whole population, with a larger number of cases coming from smaller centers; this could be explained by the stronger selection criteria used in smaller centers and could represent a potential bias. The interesting result is that most postoperative complications appear beyond 90 days, which differs greatly from major abdominal surgery, and again, demonstrates many reasons why the follow-up period for bariatric patients should be longer.

It is clearly demonstrated that the expected outcomes in patients without the described exclusion criteria should be good in at least 95% of cases.

The main question that arises is: what should the outcome benchmark be in patients with comorbidities that are currently often found in bariatric patients? Such an analysis could have been more useful to understand which patients were suitable for centralization. Based on the current study, it would appear that easy and uncomplicated patients could be operated on in a mid-volume center as well.

In summary, this important data collection and robust statistical analysis could have focused on a more complicated group of patients, which may better represent the ''real world'' that the authors cited at the beginning of the discussion.

Response From Marco Bueter (Zurich, Switzerland):

Thank you very much for the encouraging comments and underlining our finding on the temporal distribution of postbariatric complications. Indeed, our data highlight the fact that bariatric surgery requires a long-term follow-up, since the rate of reoperations and reinterventions did not reach a plateau by the end of the second postoperative year. We agree that similar studies are needed to identify the best achievable outcomes in higher risk patients as well. Our plan is to perform such studies in the future, with a special focus on the best achievable outcomes in revisional bariatric surgery. However, as a pioneering step to introduce the concept of benchmarking in bariatric surgery, we aimed to identify the best achievable outcomes in a homogenous cohort of low-risk patients. The hypothesis was that the best procedure-specific outcomes could be achieved when low-risk patients are operated on in high-volume centers by experienced surgeons. This methodology was summarized by Staiger et al in the *British Journal of Surgery* earlier this year, and has been previously applied in other fields of visceral surgery. Although the study population represented only 14% of the total bariatric caseload, the 5577 identified benchmark cases provided a meaningful cohort for the purpose of this study. Given the various possible applications of the identified benchmark cut-offs (ie, the unbiased comparison of different centers by accounting for differences in case-mix, the identification of ''out of benchmark'' cases to be presented at morbidity-mortality meetings, the validation of the introduction of a new procedure, the identification of cases suitable for the teaching of surgical trainees, etc.) we consider our study relevant for the improvement of clinical practice and stimulation of similar initiatives.

Mario Morino (Torino, Italy):

Thank you for this interesting paper and concept of benchmarking. This is quite a complex concept for surgeons. Don't you think that this concept and these results, based only on 14% of the population, might have a devastating medico-legal impact? Could you please comment on this point.

Response From Marco Bueter (Zurich, Switzerland):

Thank you for your important question. The quality control of surgical performance is getting more and more meticulous each year. Since the surgeons are at the forefront of detecting and managing postoperative complications, I believe that quality improvement initiatives should be led by the surgical community, instead of politicians, health insurance companies or other policymakers. The decision to set benchmark cut-offs at the 75th percentile of the participating centers' median outcomes supports the aim of providing achievable goals. The establishment of surgical benchmarks will need to be updated at a regular interval, reflecting the current best practices. Overall, I do not expect devastating medico-legal consequences; on the contrary, I foresee that our study will inspire quality improvement initiatives.

Bas Wijnhoven (Rotterdam, The Netherlands):

Thank you for your presentation. If you combine data from many centers, you need to have agreed definitions for complications. We need to be sure of the validity of the databases that you used, to make sure no complications are left out. So, how did you establish this, and what was your plan to make sure that all of the data entered were of good quality? Did you check them?

Response From Marco Bueter (Zurich, Switzerland):

Thank you very much for your comment. First, before we started collecting the data, we preregistered our study protocol at ClinicalTrials.gov to define endpoints of interest in a transparent fashion. Second, criteria for the participation of the centers included a prospective institutional bariatric database and at least 1 previous publication on surgical outcomes, to guarantee the sufficient quality of available data. Third, the complications were graded according to severity by using the Clavien–Dindo classification, which was emailed to every coauthor as soon as they confirmed their willingness to participate in the study. Fourth, each submitted case was controlled for completeness by the first author of the study. If an incomplete case submission was detected, the coauthors were recontacted and asked to provide additional information. Overall, in a retrospective study it is challenging to achieve complete data collection. Nevertheless, we tried to create conditions that minimize the risk of underreporting.

Norbert Senninger (Münster, Germany):

I would like to join the other speakers in applauding your study. My comment or question concerns your exclusion of patients with a BMI of above 50. We know that the sleeve gastrectomy is especially valuable in patients with a BMI above 50, for which other procedures do not work. Could you please comment on the data that relates to patients who are considered to be ''super obese''? Also, please make sure that the modern techniques of dealing with leakages are available at all of the centers. I'm mainly addressing the endo-vacuum approach, which is not of very widespread use in bariatric surgery, even though it can save lives.

Response From Marco Bueter (Zurich, Switzerland):

Thank you very much for your comment and applause, which I highly appreciate. The reason we considered patients with a BMI $>50\,kg/m^2$ as nonbenchmark cases was based on previous cohort studies (Flum et al N Engl J Med 2009), showing that the risk of perioperative complications is increased in the ''super obese.''

With regards to your second question, all included centers were high-volume academic bariatric referral centers, which were supposedly equipped with state-of-the art technology and disposed of dedicated multidisciplinary teams to detect and treat postoperative complications.

John Reynolds (Dublin, Ireland):

Congratulations on your paper. I just have a brief question. When you introduced your topic, you talked about complex operations and how they have been benchmarked. However, you based your population on the easiest cases, whereas there is a great need to benchmark the more complex metabolic surgical cases or bariatric surgery for end-stage renal surgery, for example. I'm just wondering whether you're researching this. Why did you choose the easiest cases over the more complex ones? Most benchmarking, with all of its implications, has been based on complex cases.

Response From Marco Bueter (Zurich, Switzerland):

Thank you very much for your question. This concern has been covered in part by my answer to the first discussant. When introducing a new methodology in a field, it seems intuitive to follow a previously validated approach, that is, to establish benchmark cut-offs in low-risk cases, to demonstrate best achievable outcomes in a given field. As a next step, future studies will need to focus on higher risk cases and complex clinical scenarios, in addition to the best achievable metabolic outcomes.